A Privacy-Preserving and Generalization Model for Medical Imaging Based on Deep Learning

Titinunt Kitrungrotsakul [a], Lanfen Lin [b*], Ruofeng Tong [b], Jingsong Li [a,c], and Yen-Wei Chen [d*]

[a] Research Center for Healthcare Data Science, Zhejiang Lab, China.

[b] College of Computer Science and Technology, Zhejiang Univ., China.

[c] College of Biomedical Engineering and Instrument Science, Zhejiang Univ., China.

[d] Graduate School of Information Science and Engineering, Ritsumeikan Univ., Japan.

E-mail: titinuntkitrungrotsakul@zhejianglab.com

Abstract: Currently, diagnostic tools based on artificial intelligence (AI) have shown promising results in multiple medical domains. Classification, detection, and segmentation of medical imaging are fundamental to the medical image computer-aided diagnostic system. Unfortunately, the availability of public data hinders the performance of medical image applications due to the fact that robust medical image applications require large-scale datasets. Because of these reasons, multi-centre training schemes are superior to single-centre training schemes because they can extend the scope of the existing cases, facilitate more data sharing, increase the training data, and make the training model more robust. With the existing multi-center training schemes, the lack of public data for training the AI model can be solved by collaborating and sharing the data between centers. However, the availability of datasets for AI training and evaluation is obstructed because of medical data privacy regulations and data heterogeneity. Those concerns about using clinical information to publicize the data are inappropriate. Confidential and sensitive information about patients that often cannot be distributed outside the institutions, especially when de-identification cannot be guaranteed. In this research, we propose a generalization model for medical imaging based on federated learning for multi-center training schemes. Our approach to dispatching medical image information across clients is privacy-preserving. Furthermore, we design a learning paradigm for local learning to generate domain distribution shifts and generate model generalization in medical imaging. Rather than learning the center model from data or features from clients to generate the center model, our model learns to generate a refined version model for each client based on feature distribution. Empirically, we show that our method can re-calculate the center model without asking other clients to do re-training when some clients update their data or new clients are added to the network. The effectiveness of our method is demonstrated by its superior performance over state-of-the-art methods.